Aalto University
School of Science
Master's Programme in HCID and ICT Innovation

Mario Lopez Batres

Integrating Spatial Audio in Voice Guidance Systems

Master's Thesis
Espoo, February 19, 2021

Supervisors:	Professor Mika P. Nieminen
	Professor Sandra Pauletto, KTH University
Advisor:	Dr. Francesco Grani
	PhD. Claudio Panariello

Aalto University
School of Science
Master's Programme in HCID and ICT Innovation

ABSTRACT OF
MASTER'S THESIS

Author:	Mario Lopez Batres		
Title:			
Integrating Spatial Audio in Voice Guidance Systems			
Date:	February 19, 2021	**Pages:**	56
Major:	Human-Computer Interaction and design	**Code:**	SCI3020
Supervisors:	Prof. Sandra Pauletto Prof. Mika P. Nieminen		
Advisor:	Dr. Francesco Grani PhD. Claudio Panariello		

Navigation systems are commonly used in our daily lives. Research has shown that spatial audio presents one opportunity for more effectively communicating to the driver the direction of the next manoeuvre. This thesis project proposes a new feature for the spatialisation of the audio cues triggered by a mobile navigation system by using a virtualised vector-based panning (VVBP) architecture for the encoding and decoding. The prototype developed during this thesis enables the spatialisation using headphones- or speakers-based systems. This study aims to promote a new sound experience to the user, which can be used to increase the safety and performance of driving.

Based on an expert review and a user test, the application was tested on different scenarios. The participants selected during these sessions were part of HERE Technologies, which made possible to reach design experts who knew the current application provided by the company beforehand, making easier the comparison with the proposal. This selection could also present a limitation on the study since the users might have a personal bias for seeing new features in a product which have already worked on.

Analysis of the results obtained during the testing session demonstrated high satisfaction with the feature by the users and a better understanding of their surroundings. Consequently, this indicates that spatial audio can improve the performance of driving by introducing a new source of information for positioning the next turn or obstacle. Further research is needed to identify other factors that could strengthen the effectiveness of the product.

Keywords:	spatial sound, navigation guidance, 3D-audio
Language:	English

Acknowledgements

This thesis represents the end of my double master's degree in Human-Computer Interaction and Design, and therefore, the experience that I obtained studying at Aalto University and KTH University.

Over the last two years I have lived in two wonderful countries (Finland and Sweden) was I got to know the Nordic lifestyle and I met a lot of new friends who accompanied me during the studies. I can not thanks to them enough for making the last two years something to remember.

For this reason, I gratefully acknowledge the assistance that my supervisor, Claudio Panariello, for guiding me during this process, and my examiner, Sandra Pauletto, for providing feedback based on her experience in order to create a stronger thesis report. Furthermore, I would like to express my deepest appreciation to Francesco Grani for offering me the possibility of conducting the thesis at HERE Technologies, and for being an excellent mentor, playing a decisive role on this process.

Lastly, as usual, I would like to thanks my family for their profound belief in my work and abilities and unwavering support since the beginning.

Berlin, February 19, 2021

Mario Lopez Batres

Abbreviations and Acronyms

SDK	Software Development Kit
UI	User-Interface
UX	User-Experience
HCI	Human-Computer Interaction
PoC	Prof of Concept
TDRT	Tactile Detection Response Task
ADAS	Advanced Driving assist systems
TOR	Take-Over request
HRTF	Head-Related Transfer Fucntion
VBAP	Vector-Based amplitude panning
VVBP	Virtual Vector Based Panning
ILD	Interaural Level Differences
ITD	Interaural Time Differences
AR	Augmented Reality

List of Figures

Contents

Chapter 1

Introduction

Navigation guidance is a common task in our daily lives. It can be seen how this functionality is integrated into many aspects of our routine, starting from looking into a paper map at a shopping mall to reach one specific shop, until one of the many digital options which can be found nowadays, such as, tracking a package, commuting to work, or cycling and hiking applications.

Therefore, there is a significant interest in finding a way to communicate the best navigation experience when providing a service. One approach to improve the navigation experience is to target the auditory system offered in addition to the visual channel (paper maps, signs, billboards, application UI). On this field, auditory systems are considered as a better candidate than other alternatives, such as tactile or olfactory systems due to the flexibility and scalability.

In the real world, it is unavoidable not to get distracted when conducting navigation tasks since there are many external factors which can not be controlled by the user, such as ambient noise, cognitive load when conversing or holding mental directions.

Auditory spatial cues convey high effectiveness when directing visual attention [26]. Klatzky et al. (2006) showed with their experiment the superiority and advantages of virtual sound for guidance when cognitive load is present, and the implications for guidance systems for the visually impaired users [16]. This solution is one of the many applications in which spatial sound can be implemented.

This spatialisation is mainly achieved by using two techniques [27] [29]:

1. By providing signals directly at the ears in the intended positions as indicated in Figure 1.1a
2. By controlling a sound-field using a loudspeaker setup as illustrated in Figure 1.1b

(a) Headphones setup

(b) Loudspeaker setup

Figure 1.1: Techniques for spatialisation ©Maryia Satsura

Therefore, one alternative is to make use of stereo headphones or earphone-based systems. On these devices, when virtual sound signals are displayed, the listener can perceive in different direction audio sources. Using this solution, obtaining an accurate perception of distance is difficult to achieve [17]. A second approach is to make usage of loudspeaker-based systems to spatialise the outcome sound, being applied, for example, used in the concert halls [38].

Certainly, spatial sound is already used in different sectors, but in automotive, it is only starting to be explored. For example, car-makers, such as Audi, Porsche, BMW or Renault make a severe effort in delivering well-defined sound experiences to their customers [34] [31].

Some of these companies are starting to explore spatial sound as a concept to be integrated into their products. The goal of this new target is to provide to the driver an additional source of information which could help to localise any event even when there are no visual references.

The project of this thesis is conducted at HERE Technologies, which is recognised as the number one location platform, or location-centric development environment, by Ovum Location Platform Index, basing the result on two main criteria: the completeness of its platform and the platform's market reach [39] for it innovating location solutions and ultimately enabling an autonomous world for everyone.

HERE Technologies has been working on analysing different options of integrating spatial sound into their Software Development Kit (SDK) for many years.

The scope which has been defined for this master's thesis is to continue one of the projects started by HERE Technologies a few years ago, which focused on exploring the possibility of extending the driver's senses and increasing the situational awareness around the vehicle by administrating a common sonic language of communication, in which sounds are informative concerning content and spatial characteristics.

The purpose of the thesis is to develop an Android application aiming to work as an initial prototype for understanding the feasibility of the implementation of a new potential spatial audio feature which could be included on HERE's Software Development Kit (SDK) and to test if the resulting implementation matches the users' expectations when perceiving the defined spatial audio trajectories. By enabling spatial audio, HERE Technologies expects leveraging the company's location intelligence in the audio augmented reality (AR) domain. For this reason, during this thesis was required not only to develop the prototype, but also to support the definition of the methodology which was used to test and evaluate the application, to recruit the participants for the user test to be conducted, and to gather and analyse the data regarding user's perception.

As automotive and other industries widely adopt the navigation solutions offered by HERE Technologies, the relevance and importance of the research fit into a critical change on the perceived experience of millions of users if the proposed solution is integrated.

Chapter 2

Background

Driving personal vehicles has become a daily routine in our lives. According to the department of economic analysis of the University of Zaragoza, at least 20% of the European workers spend more than 90 minutes commuting to/from work daily. It should be noted that the mobility of the European population is based on the use of private vehicles (50% use private vehicles daily while only 16% use public transport). [11]

One question that arose in the field of Human-Computer Interaction (HCI) was how to provide feedback to the driver to establish the best possible solution when indicating what is happening in the surroundings. The increment of the awareness can reduce the distraction of the driver, which could arise by different channels, such as a conversation with passengers, calls on a mobile phone, or by extensive use of in-vehicle information systems [33].

Since driving is an activity which requires eyes and hands, traditional driver-vehicles interactions techniques were recognised a long time ago as inadequate and unsafe. New investigations increasingly emerged, expecting to create new alternative methods [36]. Over these alternatives, tactile-, speech- and auditory-based systems emerged as the favourite candidates to replace the visual ones.

People can perform multiple activities at the same time, however, they are only able to accomplish these concurrent tasks as long as there are enough cognitive resources available, or there are no conflicts in their demand [37].

Speech is considered as the only wide-band communication channel which does not require significant engagement while driving. Therefore, spoken systems, which aim to enable users to interact via natural dialogues by making use of simple grammars, were one of the most attractive features for in-car systems since they allowed hands- and eyes-free interaction [35] . Nevertheless, these benefits could be forfeited if drivers need to increase the cognitive effort during this dialogue, increasing the distraction on the users and drawing their attention, commonly defined as the concentration of awareness to a specific source of information [14].

Speech-based interaction should only demand auditory perception, verbal memory, and vocal response since driving requires the usage of resources associated with visual perception, spatial working and manual response. However, traditional interaction solutions do not follow this principle, which is why other authors have not found a solution to using these systems. For example, Vetek and Lemmelä (2011), after investigating how different dialogue strategies affect the cognitive load and the driving performance, concluded that secondary speech tasks have a significant adverse effect on the performance of driving [36]. Kun et al. (2007) found that users tend to drive more dangerously when having troubles using these systems [18]. Redelmeier and Tibshirani (1997) also determined that hands-free interaction while having calls has not proven to increase the safety of the driver [30].

On the other hand, voice control systems seem to be less distracting since visual-manual interfaces take the drivers' attention away from the primary task of driving when involving dialogue. As listening is more automated than speaking, speech-based systems require a higher cognitive workload than listening. Accordingly, these systems present longer Tactile Detection Response Task (TDRT) response times and miss rates. Hence, interface designers should be aware of this fact when creating interactions and should use simple speech commands that have limited information complexity [3].

Another alternative is tactile interfaces, such as haptic feedback. Tactile interfaces present some advantages over visual or auditory systems since they do not distract the users as much as the alternatives. Nevertheless, they are commonly used for short notifications and not for communicating any complex messages [32].

Meanwhile, auditory user interfaces are a good alternative for being flexible and scalable. These interfaces present different approaches, starting from simple non-speech cues, to earcons, auditory icons, synthetic speech outputs or audio representation of multivariate and multi-dimensional data compounds. Moreover, auditory interfaces do not need visual information and can capture the user's attention even if the source is hidden from view. Sodnik et al. (2008) compared the effectiveness and efficiency of two new auditory interfaces to a standard visual interface [32]. This comparison aimed to evaluate driving performance, perceived workload, task completion, and overall user satisfaction. Even though the three interfaces presented a similar task completion time, the results also concluded that by using the auditory ones, not only was the driving performance significantly better but also a lower workload was perceived when using the auditory interfaces leading to an increment of their sense of safety and lowered their stress levels.

With the rise of automation and in-vehicle entertainment and information Advanced Driving Assist Systems (ADAS) over the last decades, researchers started studying how would they influence on the driving takeovers. Damböck et al. (2012) stated that highly automated driving could reduce the number of human failures by increasing the safety of the diver [7]. Borojeni et al. (2016) proposed the use of ambient Take-Over Requests (TOR) which focus on the peripheral vision of a driver to reduce the visual processing demand and time restriction that driving tasks present. By lowering their visual processing demand, it is possible to help drivers to relocate their attention from a secondary task proving that ambient displays led to a shorter reaction time without increasing the workload and increasing their comfort [2].

A specific field that has been explored over the last years is how to communicate this attitude through audio cues. Once mono cues were established in-vehicle systems, researchers started to think about spatial sound, which can represent audio items in different locations providing extra information to the user.

Spatial sound has been researched by many authors over the last years concluding that, even if humans can perceive directions of sounds sources, generally named as directions of auditory objects, localising a sound presents a challenge for humans. Localising a sound is a challenge to our nervous system. When describing the visual system, this task is focused by the optics of the eye and maps of visual space are found in our central nervous system. This approach is not feasible when describing sound since the sound wave generated is diffracted by its interaction in head and ears. While it is easy to locate objects when looking at visual references, the sound produced by an external source is not. [24]

The spatial perception of sound is based on the interpretation of several cues. These cues are extracted from the surroundings of the listener. When using artificial sound fields, the perceived direction and distance of the sound might not match with the original ones [22].

Electronic systems can artificially position sound in the space. This positioning task can follow one of two approaches. These systems can generate:

- Phantom images by amplitude profiling the feeds to multiple speakers aiming to create immersive auditory displays by applying a sound signal to loudspeakers with different amplitudes

- Try to mimic the complex changes in timbre, delay and amplitude which occurs directly at our ears based on the Head-Related Transfer Function (HRTF).

One generic method for virtual source positioning is Vector-Based Amplitude Panning (VBAP), which is an amplitude panning method for positioning virtual sound sources in three-dimensional arbitrary loudspeaker setups. In early studies, it has been proved that, if the loudspeakers are placed symmetrically with the median plane of the listener, the positions can be controlled with good accuracy [28].

Presentation over headphones can be very pleasant, approaching the listener's capabilities for localisation. A common problem when using binaural recordings is that the well known front-back reversals might occur due to the lack of head rotation, which is used by humans to distinguish the front from rear sound sources [23]. Magnitude differences between our ear canal signals are named the interaural level differences (ILD) and the temporal differences or interaural time differences (ITD) [10].

While the field of spatial sound has been investigated, this domain is unexplored when describing an automotive environment. However, it has been confirmed that traditional driver-vehicle interactions are not appropriate, and some benefits of leveraging speech, such as cognitive effort during the dialogue interaction have been established [36]. Indeed, in-vehicle voice interactions seem to offer a less distracting way to interact with the driver, enabling a better UX while driving by providing a more secure procedure [3].

2.1 Previous work

Early stages

In the past, HERE Technologies' team started understanding the potential advantages of integrating spatial audio guidance during the navigation. The experiment carried out consisted of observing how the driver reacted while driving. This section provides a brief introduction to the stages carried out by the UX team in the past intending to show how the project was established.

The background for these stages was to establish a "sonic language", or a language to communicate with the user through sounds. These sounds were informative for their content and spatial characteristics on a navigation guidance system. Sound cues were designed and validated on two different configurations: an ambisonics setup using a full-sphere surround format, and a binaural setup recording the sound using two microphones. Both mounted inside of an electric car intending the catch the direction of the proposed spatialisation experiment.[12]

During the experiment, there was not a real-time application which triggered the audio cues within a spatial sound. Instead, some routes were simulated by rendering audio files and replicating a situation as if the navigation system would be working in real-time triggering the audio files in the right moment.

Figures 2.1a and 2.1b shows the setup of both approaches:

(a) Binaural and Ambisonic setup

(b) In-vehicle setup

Figure 2.1: Previous setup ©HERE Technologies

To extend the driver's experience, the increment of the situational awareness around the vehicle was required. For this reason, the common sonic defined three stages [12]:

1. Catch: Catching a driver's attention without being invasive.

2. Inform: Informing the driver on the nature of the occurring event.

3. Aid: Aiding in an action to be taken which ideally would support the drive towards an action that should be taken,

During this stage, two preliminary use cases were chosen for the study [12]:

- An ambulance coming from behind the car
- A cyclist located in a blind spot on the left side of the vehicle

The results of the experiment concluded with highlight promising results. The drivers showed satisfaction with the proposed feature, spotted the spatial location correctly and commented about having the feeling of being in control of the situation. As expected, the Ambisonic setup presented a better map for those cues positioned in the front of the driver. Nevertheless, the binaural made a more private and personal experience, creating even more aware of a dedicated message. [12]

After analysing the data gathered, the team concluded that there was an opportunity to keep working on a new solution including spatial audio and the team decided to continue with the project and to prototype a real product for further tests. Therefore, the need for the proposed application in this thesis emerged.

2.2 Problem and objective

Synthesised voice instructions in the field of guidance navigation are widespread. A standard solution, since the driver is making use of his visual channel while driving, is the usage of spoken instructions to deliver information regarding the upcoming manoeuvres required to reach a destination.

While monaural solutions (Figure 2.2a) are sufficient to provide content, over the last years, several improvements have been made in the field of stereo (Figure 2.2b) and spatial sound reproduction (Figure 2.2c). By adding more audio channels, an enhanced immersive audio experience is created where sounds can flow bidirectional audible perspective in the case of stereo, or in three-dimensional virtual space in the case of spatial sound.

Including the location context, a new sound information source is integrated, conveying rich auditory information such as position, movement and distance [12]. Nevertheless, these new solutions are not generally implemented in navigation systems. Therefore, the problem to solve is how to fulfil the context-awareness by providing spatially relevant auditory information.

(a) Monaural (b) Stereo (c) Spatial

Figure 2.2: Monaural vs spatial sound ©Maryia Satsura

The purpose of this thesis is to develop a prototype to spatialise the audio cues triggered by HERE Technologies' SDK by integrating an external library to render the spatialisation of the audio cues. The prototype aims to make a proposal for the implementation of the new functionality, allowing leveraging HERE's location intelligence and HERE's specific IP in the AudioAR domain. Furthermore, during the thesis was designed, prepared and executed a user test to gather information regarding the user's perception of the spatialisation in order to understand the level of acceptance of the proposal, and therefore, to validate that the proposed solution was correctly perceived by the user.

Through this new functionality, in the future, HERE Technologies intends to find a connection of improvement of the observed driver's point of view, such as safety, time-response and performance, and the integration of spatial sound cues in its application. Nevertheless, this was set out the scope of the thesis.

2.3 Research questions

Before designing the research questions, it was necessary to understand which data was provided by HERE Technologies' SDK during the navigation guidance, and how could be possible to implement the new spatial audio feature within the aforementioned data. Therefore, two preliminary steps were defined:

- What data is provided by HERE Technologies' SDK?
- How can the audio cues triggered be spatialised within the provided current data?

Once the previous prerequisites were positively answered, it was possible to confirm the feasibility of the spatialisation of the audio cues triggered by HERE Technologies' SDK. Therefore set of research questions were established intending to collect information from the user's perspective about the designed spatialisation. For this reason, the RQs focused on one of the three aspects of usability, the satisfaction [1].

Since HERE Technologies focuses on delivering high and premium quality and aims to build a loyal customer base, the objective of this project emphasised on this aspect. Therefore, two research questions (RQs) were defined:

RQ1. How does the user perceive the condition of the spatialisation of the audio cues?

RQ1. How does the spatialisation of the audio cues fulfil the user expectations in regards to the perception of the spatial audio?

The first research question focused on the quality of the sound. HERE Technologies tries to deliver a new experience to the driver. Hence, the sound offered by the prototype must be clear, have a constant level of perceived volume and an implementation which allows a continuous movement of spatialisation. On the other hand, the company provides a solution which is welcome by the user, and therefore it is mandatory to create a prototype based on a user-centred design which reflects the needs and expectations of the driver. Thus, the second RQ addresses this point of view establishing his expectations as the main point of view of the designed proposal.

By addressing the research questions, the company expected to gather enough data from the user to optimise the experience perceived, and the trajectories defined for the spatialisation of the audio cues.

Chapter 3

Methods

Until just a few years ago, the availability of cross-platform, lightweight, and good-performing spatial audio libraries to be integrated into mobile applications without a significant effort was very limited. Therefore, the project started back in the past was parked aside until new potential libraries were published.

The thesis started once different development and improvements on the spatial audio libraries regarding size and performance were noticed. In the following subsections are discussed the stages which were conveyed in order to conduct this thesis.

3.1 Procedure

This document makes use of the 5-stages design thinking model proposed by Hasso-Plattner Institute of Design at Stanford (dschool) to give a better understanding of the methodology which has been used during this project [6]. The stages in which the method is divided are the ones described below:

1. Empathise: The first stage of the method consists on consulting experts or engaging and empathising with people. It aims to understand the potentials users and their needs in the real world.

2. Define: The second stage, makes use of the needs determined in the previous stage to identify and to define a statement of the problem. The problem should be analysed in a human-centred manner.

3. Ideate: After the information has been gathered, and having a solid background, it is time to generate options or alternatives ways of viewing the problem. Once the ideation session has finished, a selection between them should be made to investigate and test the best ideas.

4. Prototype: By the end of this stage, a simple experience through the previous options are created. The prototype can be expanded as long as it is required.

5. Test: Evaluators rigorously examine the best solutions defined during the stage of prototyping.

Design thinking methodology is considered as a non-linear method. Therefore, the order of the stages might vary. Besides, they can be repeated as many times as is necessary since the process is iterative. The common factor in deciding how many times to iterate the method is usually the budget of the project. Figure 3.1 illustrates how the iteration between stages is commonly done.

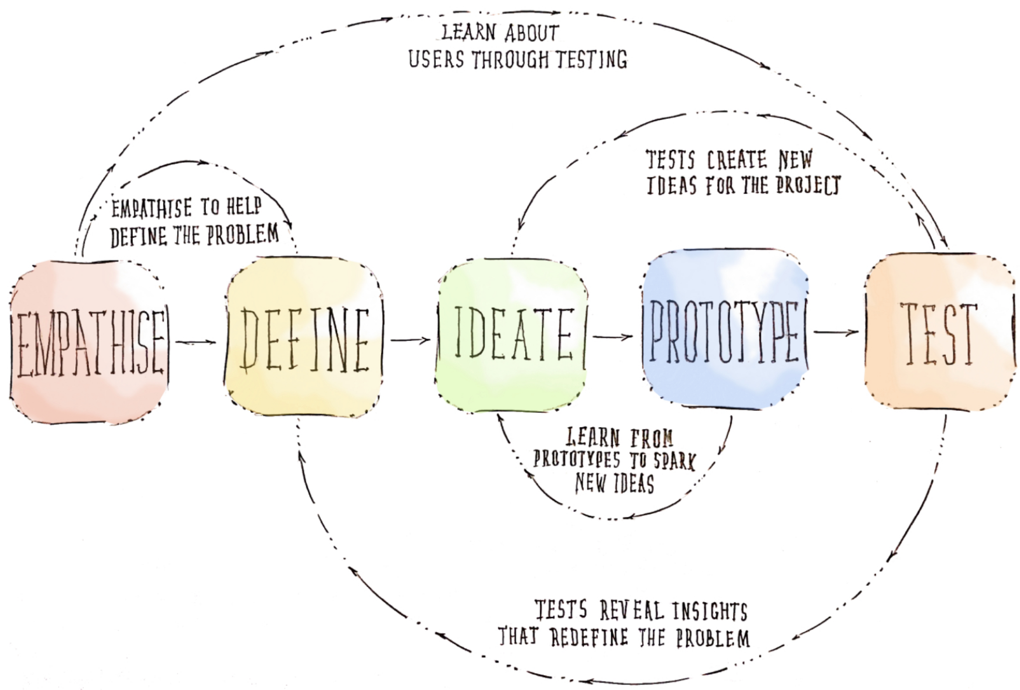


Figure 3.1: Design thinking: A non-linear process

Since this thesis is the continuation of the project above-mentioned in section 2.1, the stages to be conducted were: Ideate, Prototype and Test. The stages were conducted by two members of the HERE Technologies' UX team, including the author of this document.

The task of developing the prototype was carried out by the author of this thesis. Nevertheless, the task of designing and defining the spatial trajectories to be used to achieve the spatialisation of the audio cues was a collaborative assignment between the author and an audio expert part of HERE Technologies' UX team.

Below is presented an introduction to the tasks which were conducted during the project reported in this document.

3.1.1 Ideate stage

At the beginning of the thesis, it was required to define a set of minimal features which should be present in the prototype, as well as the pros and cons of different frameworks which could be utilised to develop the prototype.

For this reason, the technique of Brainstorming was used during one ideation session between the two members of the team, including myself, involved in this project. The session was not recorded since it was conducted as an informal discussion between the two UX experts.

3.1.1.1 Minimal features

Once the ideation session was concluded, and the data generated was analysed and discussed by the same two members implicated on the Brainstorm session, a set of minimal features which should be included in the prototype were generated. The ideas were divided into two different categories: the features to be supported by HERE Technologies' SDK, and the ones supported by a third party library. The requirements which were defined are listed below:

1. HERE Technologies' SDK: Three main requisites were established:

   R1. The prototype needs to have a navigation system integrated to allow calculating routes to guide the driver from A to B.

   R2. The prototype's guidance needs to work as live-navigation tracking your current position in the real-world, and as a simulation by replicating the guidance as the user was moving

   R3. The prototype requires an engine which allows transforming strings, in this case, the audio cues, into audio buffers or audio files.

   R4. The prototype needs to be able to spatialise the coming audio cues triggered by the SDK to provide a more robust and immersive experience.

2. Spatial audio characteristic

   R1. The selected external library had to be able to spatialise along a circular trajectory and not only a linear one

3.1.2 Prototype stage

The framework chosen as a tool on this project was Android Studio in order to develop the mobile application. Moreover, the prototype had to make use of an external library to spatialise the audio cues triggered by HERE Technologies' SDK. The mobile application would help to easily test the new feature as a pedestrian or connected to a car while driving.

Since the sound feature quality of this prototype was more crucial than the layout or visual impact of it, there was no need to create a low-fidelity prototype, but only a final version of it which included the audio functionality. For this reason, no mock-ups were created during the execution of the project established on this thesis. Furthermore, the prototype did not follow any of the visual standards of the company.

The prototyping stage was conducted along the whole project implementing new features which came up during the process of the expert review.

3.1.3 Test stage

To review and test the prototype, two methods were defined: a UX expert review, and a user test. The UX expert review was conducted in order to validate that the outcome audio and the trajectories which were spatialised before testing it on users.

3.1.3.1 UX Expert Review

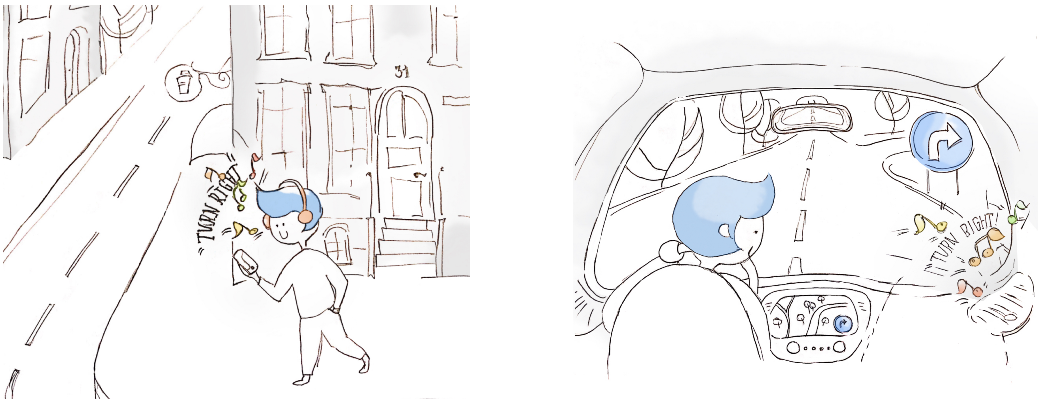
The expert review was conducted by two experts of the HERE UX Team directly implicated on the project. This revision was executed at different development stages of the prototype and therefore evaluated different scenarios. The setups which were used for this method utilised different audio equipment: headphones as shown in Figure 3.2a and in-vehicle systems as illustrated in Figure 3.2b. Inside the vehicle, the connection was established via Bluetooth, granting an easy way to connect the audio outcome from the phone to the speakers of the vehicle, and covering the most realistic scenario.

3.1.3.2 User test

Regarding the user test, the tests were carried out only wearing headphones completely remotely using the simulation mode of the prototype.

Since the testing stage was based on a user test, it was required to recruit participants. Due to the COVID-19 pandemic, the original plan to conduct this user test had to be changed. The social distance imposed in 2020 concluded with the closure of some buildings at HERE Technologies, making it not possible to access to the simulator owned by the company. Therefore, a new way to perform the test remotely was defined. During this method, users did not navigate using the system, but tested the prototype from their working spaces or homes.

The test exposed two pre-defined routes in order to ensure that the user heard at least once all the spatial trajectories. Once the user finished listening to the audio guidance presented, he was granted access to an online survey. The survey focused on obtaining qualitative data regarding the user's satisfaction with the product, some quantitative data regarding the spatial trajectories, and information about their expectations and expected results.



(a) Headphones

(b) In-vehicle

Figure 3.2: Approaches for testing ©Maryia Satsura

3.2 The prototype

The resulting functional prototype implemented some of the features offered by HERE's SDK and the additional support of a spatial audio library working, in the end, as a fully functional Android application. For the intended purpose of the application, which is mostly an audio-related prototype, as mentioned above, the UI of the prototype was not aligned with the HERE brand and design style guidelines.

To achieve the goals mentioned in section 3.1.1.1, the prototype was split into two major components:

1. Navigation guidance system: The first component required was a navigation system. HERE's SDK allowed to build an immersive native app

which was able to enable routing, search functionality and real-time guidance.

2. External library: The second component was a spatial audio rendering engine capable of receiving as input a monophonic audio file and coordinates to describe where the sound should be placed in time and space. The library presented an alternative to ambisonic for a spatial audio containment, format and playback without any abstraction layers utilising a virtualised vector-based panning (VVBP) architecture for encoding and decoding processes, which made it the primary candidate.

Lastly, in order to describe the actual prototype, this document splits the description into three dimensions:

- Functional fidelity: The prototype was built using Android as a framework and therefore it presented fully interactivity. The final dynamic prototype used the real-data triggered by HERE Technologies' SDK as any other application in the real market would do. Therefore, no hard-code or mock data was generated.

- Visual fidelity: Since the prototype was created using Android, all the components which were used for its implementation followed every standard. Nevertheless, it is important to mention that the prototype did not intend to mimic any other product of the company, and therefore could not be compared.

- Feature fidelity: A final dimension is the spectrum of complexity completeness. The resulting prototype fully implemented the features which were required, such as search engine, routing, navigated and simulated guidance. Nevertheless, while all the already mentioned functionalities were implemented, some potential limitations were noticed:

  - The prototype used the 3.15 version. Having the last version of the SDK would have created a more realistic implementation of the behaviour of the current navigation system offered by HERE Technologies.

  - The prototype could set a route by selecting an initial and destination address by using your current location or the name of a street. Nevertheless, for example, it was not possible to search for places. This factor should not really affect the user's perception since the scenario which was given contained a predefined route selection

3.2.1 Hardware compatibility

Two hardware specifications were defined as the requisites established to validate the implementation of the prototype. The ideal prototype had to perform not only using headphones but also with in-vehicle systems. Accordingly, there were two connections to be considered:

1. The first one was the usage of a phone and its earphone jack, and the usage of the Bluetooth connection.

2. The second setup would allow us to make use of the prototype on different devices, such as the helmet of a motorbike, headphones/earphones or the speakers of the car.

In the end, both configurations could be implemented by a unique solution.

3.2.2 Structure of the prototype

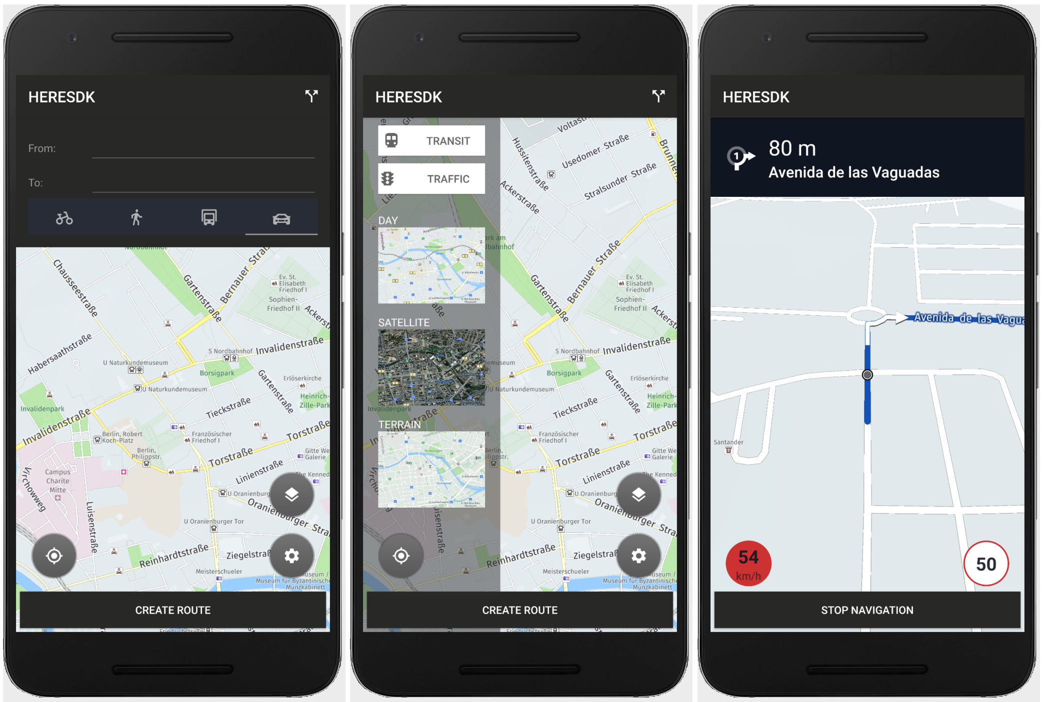
The structure of the application was divided into two main functionalities. The navigation system, which was presented as the user interface (UI), and the elements of the external library stayed in the background spatialising the audio cues triggered by the SDK.

The prototype showed to the user a map implemented using HERE's SDK meanwhile, the external library was running in the background to spatialise the audio triggered by the SDK each time it was needed.

Regarding the UI, the prototype mainly required two main screens. The first one was an activity which intended to prepare the navigation, meaning its settings, showed as floating buttons in Figure 3.3b or defining the initial and destination addresses (Figure 3.3a). The second one visually guided the user during the navigation. This activity guided the user in real-time through the navigation. The reason for choosing a minimalist design was to avoid visually catching the attention of the user rather than listening to the spatial audio cues.

3.2.3 Dynamic functionalities

All the minimal functionalities that were planned for this prototype were successfully implemented. As mentioned in the previous sections, it was necessary to integrate some of the functionalities from HERE's SDK to gather the required data. The list of these features are mentioned below:



(a) Main layout (b) Schemes layout (c) Settings layout

Figure 3.3: Layout of the application (1)

- Set a route: Allowed the user to set the initial and destination address to create a route between them (Figure 3.3a).
- Search engine: Allowed the user to introduce an address and transform it into coordinates which the SDK can understand (Figure 3.3a).
- Guidance: Voice guidance during navigation or simulation of a route.
- Localisation: Navigation considering the current position of the driver/pedestrian).
- Real-time positioning: Retrieved the real positioning of the user in real-time.
- Simulation of the navigation for testing purposes: Allowed the developer to test the application without leaving the working space (Figure 3.3c).

Once the basic functionalities were running in the prototype, the integration of the external library was executed, and the main feature was implemented: Spatial audio.

- Spatialisation: Spatialise the mono audio files within a linear or circular trajectory.

Furthermore, other extra properties discussed during the development phase were implemented. Despite not being defined as essentials during the first session, the functionalities listed below allowed to provide a better experience during the testing stage.

- Allow setting a different audio threshold and spatialisation duration: The feature was implemented to allow the experts to easily tweak the values of the spatialisation.
- Integration of HERE MSDKUI: This library provided a more realistic layout for the application.
  - Enabling/disabling the speed warnings: Speed warnings could be enabled or disabled during tests providing the capability to the prototype to facilitate the user or expert to only focus on the quality of the spatialisation to avoid distractions.
  - Create a layout for guidance to show the distance to the next manoeuvre or icons to show the next turn (Figure 3.3c).
- Set a different scheme: The extra scheme maps which were implemented were satellite and terrain (Figure 3.3b)
- Add the capability of showing the traffic and public transport transit of the road (Figure 3.3b).
- Allow setting routes depending on the mode of transport: bike, pedestrian, truck and car (Figure 3.3a).

3.3 Research Methods

Before conducting the ideation session to define a set of minimal requirements, it was needed to generate potential use cases could the spatial audio feature be used in the future. For this purpose, Personas and scenarios were designed.

Furthermore, as mentioned in 3.1.3, two methods were used to review the audio and the spatial trajectories before conducting the user test. By crossing the result of both evaluations, expert review and user test was aimed to find a connection between the proposed solution built by the experts, and the expectations of the user. Lastly, finding the same outcome on the results of both methods, the reliability and validity of the results would be proved.

3.3.1 Personas

Visualising and communicating user needs or their requirements is a common problem when describing the user-centred design, and they often disappear during the development. Therefore, these needs are not taking into account in the final system.

A persona is a tool, which attempts to bridge this gap. Personas are described as fictitious characters based on knowledge of real users and research that represents the demands or urgency of larger groups of users when taking into consideration their intentions and personal characteristics [4].

This thesis follows the indications introduced by Gudjonsdottir [13], who explains that even if personas are usually received positively and as an interesting topic by developers, personal details are questioned. Sometimes developers wonder whether the personas are based on statistically valid data. The author clarifies that personas are generally accepted without any criticism or any question. This acceptance makes the author believe that the material is rarely utilised as much as intended. The reason is that some developers claim that they have a good view and knowledge of the users and their needs.

Gudjonsdottir establishes a few indications when documenting personas. While personas are commonly accepted, developers usually identify unnecessary details which are too personal, such as "having a cat". Another suggestion is to use illustrations rather than photographs. Long and Frank determined that illustrations reduce the effectiveness, but she believes that they emphasise the fact that the personas are fictitious representations of real users. Furthermore, illustrations provide more personification and avoid the situation of using a picture that could be known by someone from the team [20].

Below, some of the personas and scenarios which were discussed during the execution of this thesis are listed. If the reader wants to get more detailed information, the template used during the project can be found in appendix A. In the next section, a summary of the personas is shown.

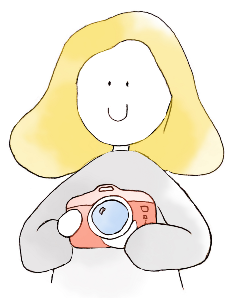


Figure 3.4: Persona 1
©Maryia Satsura

Anne Jones — 32 years old

Anne was born and raised in New Haven (USA), but she has recently moved to New York to start working at her new company. In her free time, her biggest hobby is to travel looking for new small towns to visit around her country. She believes that the only way to get to know a city is to check even the smallest street, and a car does not give her this possibility while navigating in the city centres. Furthermore, her motorbike offers the opportunity of stopping by wherever she wants without looking for a parking spot.

Jose Ramon Hinchado — 43 years old

Jose Ramon is from Madrid (Spain). He works as a taxi driver. Since Madrid is a big city, he has the opportunity to reach many new customers which makes his job more profitable.

He is happy with the conditions of his company. One of them is to continuously run the phone application that they offer, which is also the way to establish contact with the customers. This application rates each of the drivers based on a review system evaluated by the passengers so he is always looking for the most comfortable experience for them during the drives.

Figure 3.5: Persona 2
©Maryia Satsura

Nicky Doe — 18 years old

Nicky is a teenager who lives in London (England). She is a good student and loves anything related to technology and its applications. Nicky, like many other people in the world, has dyslexia, but it does not stop her from doing anything. She is about to finish high school and to start studying engineering, but first, she needs to check which university fulfil her expectations.

Figure 3.6: Persona 3
©Maryia Satsura

Moritz Müller — 29 years old

Moritz is originally from Dresden (Germany). He has always worked at a construction site, but three months ago he became a father of two twins. Since the first moment, he had clar-

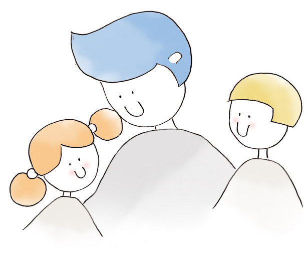


Figure 3.7: Persona 4
©Maryia Satsura

ified that he wanted to spend as much time as possible with them, so he decided to stop working during a couple of years to spend more time with his kids.

3.3.2 Scenarios

A scenario is a second tool which complements the persona by describing an activity in which the persona fulfils one of the goals [5]. It aims to build a story of the intended user and how the system could support. Scenarios should describe the usage of the system during the full duration of the activity in detail and are often illustrated. This section documents some of the scenarios which were created during the project.

Riding a motorcycle

Anne is visiting Manhattan for the first time. As always, she has taken her motorbike to reach the place. She is lost since New York is a big city, so she is using a navigation system on her phone connected to her helmet via Bluetooth to go there. Between the bike and the loud noise that comes from the streets, Anne is not able to hear clearly the indications of the system. She decided to make use of the spatial sound mode that the application offers.

Even though the volume is still low, the direction of the sound which comes only from the side that she needs to turn to helps her identify where to go even without listening to the full audio cue.

Comfort to passengers

Jose Ramon has picked up a family composed of four members in Madrid centre to go to the airport. Three of them go in the backside and the fourth one in the co-pilot seat. The family seems to have an exciting conversation, and Jose Ramon does not want to disturb them with the indications of his navigation system. Accordingly, he has connected his earphones to his phone via Bluetooth to get the instructions privately. Furthermore, the spatial mode that offers the application helps him the direction of the turn even when the family is loudly speaking.

Pedestrian

Nicky is finishing his last year at high school, and she is starting to look for a university in his city. She would like to begin his bachelor's degree in computer science, and she would like the potential alternatives before selecting one.

Nicky has decided to start looking at the University College of London (UCL), but he has never gone before. As usually, Nicky carries her phone, so she has set the address on the university on the navigation system. Once she knows the route, she takes public transport to go there.

Once Nicky arrives at the university, she starts exploring the campus. Nicky likes to walks without looking or carrying her phone in her hands not to miss any detail of the campus, so she connects the earphones into her phone. Nicky has dyslexia, and therefore she misunderstands sometimes left with right and turns the spatial mode that the application offers to get an additional source of information. With the phone in her pocket, the system helps her to know which direction to turn to by indicating it also by reproducing the sound only from one side (left/right), and she can enjoy her walk without wondering where to go.

Noisy background

Moritz is going for the first time to bring his two years old twins to the kindergarten. Moritz and his wife have chosen one in the suburbs of Dresden because a friend recommended him. Accordingly, Moritz sets the destination address in his navigation system.

One of the twins starts crying what makes the other to begin as well. Moritz is trying to focus on the indications that navigation systems are giving, but it is challenging for him to do it since the twins are crying each time louder, and it is straightforward to get distracted. In the end, Moritz turns on the spatial sound mode offered by the application, so even if his head is on driving and paying attention to the twins, he can quickly identify in which direction to turn to by determining in which ear the sound is coming from, what makes for him easier to reach his destination.

3.3.3 UX Expert Review

Expert evaluations are a cheap and effective technique to assess the usability [25] or user experience of the system [19]. Expert evaluations have been criticised by previous research on usability, pointing to the inadequacy between the expert evaluation and the problems reported by users. Nevertheless, on this project, the intention is to gather information about the spatialisation trajectories defined based on a UX expert review of two members of the UX to cross the results with the perception of the user's information gathered during the user test. Therefore, the method has been determined as adequate for this project.

The expert evaluation was divided into three different phases, depending on the maturity of the prototype. These stages aimed to define the trajectories of spatialisation, to evaluate the prototype while wearing headphones, and to evaluate it when connected to the speaker system of a car. The process intends to increase the reliability of its results by analysing the data gathered by two experts when crossing the results obtained by each of them. Below, a more detailed description of the phases is introduced. The phases are explained below:

3.3.3.1 Phase 1: Definition of the trajectories

The first phase of the expert evaluation was conducted in an early stage of the prototype. At that time, the prototype was based on a simple implementation of an example provided by the external library. During this stage, the prototype spatialised different strings triggered when clicking on the views of some buttons. As mentioned above, the phase aimed to validate the concept and define the potential trajectories which could be used in the final version of the prototype. Therefore, the implementation of HERE's SDK was not required at this stage.

At the end of the stage, the potential spatial trajectories were defined as the experts considered that was valuable and reasonable. Due to the privacy policy of HERE Technologies, the trajectories defined in this prototype are not mentioned in the document.

3.3.3.2 Phase 2: Evaluation of the prototype using headphones

The second phase of the evaluation was conducted once the prototype had the capability of defining a route and conveying the audio guidance during the simulation or navigation. Furthermore, the trajectories defined on the previous stage were integrated into the prototype to spatialise the audio cues triggered by the SDK. During this phase, the observations which were made focused on the coherence between the audio cue and the spatial movement of the sound defined.

Since the prototype allowed not only real-time navigation but also simulating the navigation of a route, it was possible for the experts to review the system when a new feature was added without leaving the working space. Regarding the setup of this phase, a smartphone running the application connected to a pair of headphones via Bluetooth was needed. Figure 3.8a shows the setup used during this stage.

Furthermore, a second technique used during this stage was immersion. In 2003, Jordan proposed a concept where the investigators use the system in a real context and evaluate it [15]. Consequently, during the expert review, a route was set in the navigation system meanwhile walking around the street. This contextual environment helped to validate the tasks proposed during the creation of personas and scenarios to be reviewed in a more realistic scenario.

The figure 3.8b presented below shows the setup conducted while wearing a pair of open ear audio glasses.

(a) Simulation in working place

(b) Real-time navigation guidance

Figure 3.8: Expert evaluation

3.3.3.3 Phase 3: In-vehicle session

Once the prototype was reviewed and validated using headphones, the last stage of the evaluation was conveyed in a vehicle. The phase helped to confirm the most critical use case of a navigation system. This phase was conducted on two different stages of the prototype.

Since the user test was conducted wearing headphones, but not in a vehicle, during this stage, a third UX expert was invited to assist to the session to review the new feature by somebody external to the project. During each session, a different external expert was invited. By inviting external experts to the sessions, the validity and reliability of the results were intended to be increased.

During this stage, the need for a more sophisticated and complex setup was necessary:

- The chosen car for this evaluation was an electric vehicle owned by the HERE UX Team. The car has been modified by adding an extra screen to mirror the smartphone running the application.

- The smartphone was connected to a pair of open ear audio glasses to be able to record the direction of the sounds. On them, it was reproduced the audio output triggered by the navigation system on the ear it was meant to be reproduced.
- A binaural microphone was placed between the driver and the passenger to record sound with a driver perspective. The record was used to analyse the data afterwards.
- The pair of audio frames were placed on the binaural microphone to the spatial audio cues gets recorded naturally superimposed to the driving soundscape. The placement was possible by using gaffa tape to mock-up a 'nose' to secure the glasses upon.
- The binaural microphone also sent the audio to an additional pair of headphones. Those headphones were used by one of the members of the UX team to parallel validate that the audio cues were correctly displayed.
- A camera was placed behind the driver seat to record the video session. Recording the session helped the experts to review the results in context.

During the evaluations, 3 members of the team were present (driving, recording and preparing the application). Figure 3.9 and 3.10 figure shows set up on this stage.

(a) Application setup

(b) Binaural microphone

Figure 3.9: Hardware in-vehicle

Furthermore, one last property that needed to be validated was the capability of reproducing the spatialisation using the sound system of a car. For this reason, the application was also evaluated by connecting the smartphone

(a) Screen and microphone

Figure 3.10: In vehicle video shooting session

via Bluetooth. Video and audio were recorded to analyse the recordings once the sessions were over. For this purpose, a video was create withing the images that were taken during the sessions and the spatial audio which was recorded by the binaural microphone.

3.3.4 User test

3.3.4.1 Participant recruitment

The participants recruited for this project were employees at HERE Technologies who were part of the design team of the company. There was not ethnic, demographic or gender/age filter for the selection. The recruitment of the participants intended to get valid feedback in a short time without breaking the data privacy policy of the company. The factor that the users have worked for years on HERE Technologies' products was defined as a limitation since it could create a personal bias for this feature on their side.

Furthermore, since all of them were experts on validating the performance of the user experience (UX), the feedback was considered more robust than the one provided by a few general users. Nevertheless, it was also seen as a potential limitation since often users and experts do not have the same expectations of a product.

In the end, 3 male users in the age between 30 and 45 years old, participated in the experiment.

3.3.4.2 Set up

The user test was conducted fully remotely. Before running the test, two documents were sent to the users containing the aim of the research and all the instruction to be followed and a link to download the APK of the android application was distributed among the participants. The APK contained the last version of the prototype, and therefore, the spatial sound functionality. In general, the task to be carried out was the simulation of some pre-defined routes in the given prototype to evaluate the spatial trajectories which were defined. These routes should be simulated as navigating in a car, even though the user could be at home since the prototype offered that feature. Below, a more detailed description of the set up:

Once the users had downloaded the APK and installed on their phones, one scenario was presented. Since the user test was conveyed online, the instructions were given as text and image, intending to minimise potential misunderstandings and increasing the simplicity of them to be followed.

In order to assure that all the users experienced at least once all the trajectories, a couple of routes were established for the user test. These routes were carefully selected and tested beforehand to ensure that all the trajectories were displayed during the audio guidance.

The estimation of completion of the test was 10 minutes. Despite the questions of the survey addressing the topics to cover, a potential gap of feedback could emerge since it was conducted completely online.

As the user test was not defined to test the usability of the UI, but to understand how the user-perceived the spatial audio, the creation of a simple scenario within a few steps was required. J.S. Dumas and J.C. Redish. defined scenarios as "situations in which the task is embedded in a reasonable and concise story" [8]. The scenario was based on a simple task which ended up in audio guidance.

Bellow, the scenario given to the participants during the user test. The scenario was executed two times using different starting and destination addresses.

1. Open the application
2. Press on the icon located on the top-right corner to show the search bars
3. Press on the "from" search bar. Once the view of "Your location", press on it.
4. Press on the "to" search bar and introduce "Invalidenstrasse 116 Berlin".

Please, note that the search engine of the prototype only allows to look for street, and not for places.

5. Select the style of navigation. Select the icon which represents a car.
6. Once both addresses have been selected, press "Create Route" button displayed at the bottom of the screen.
7. A dialogue box will be prompted. Select "Simulation" if you intend to simulate the route navigation guidance.
8. Listen carefully all the audio cues triggered during the guidance.

As the last step, two out the three users were interviewed to gather a more robust and direct feedback. Even if the interviews were not recorded, they provided a new source for gathering information where it was easy to get the thoughts and feelings of the user during the interview.

3.3.4.3 Survey

Once the users finished the task described in the user test, a link to the online survey was sent. The survey was composed of 23 questions, structurally organized into independent clusters. These clusters of questions were shown to the user depending on how the previous questions were answered. The tool which was used to create the survey was Microsoft SharePoint enabling the owner to create questions, have multiple people respond, and see a summary of the results to facilitate analysis of the data gathered.

Due to the reduced number of users, the questions gathered qualitative information about if the user observed any change in comparison with the current audio guidance that HERE Technologies provides, and if so, how did they rate the experience perceived. Furthermore, some questions focused on the spatial trajectories defined to spatialise the audio; how the movement was felt, if it was adequate or irrelevant, and if the user thought that one trajectory had to be modified, removed or added.

3.3.4.3.1 Structure of the survey

The survey was structured into a different block of questions. Each of these blocks intended to gather information focusing on different aspects of the research. Below a description of the blocks is listed:

1. The first block gathered the answers regarding the audio setup which was used. This block of question remained really simple but could be useful for finding potential incompatibilities within the hardware which was used during the test.

2. The second one checked if the user actually noticed the spatialisation of the audio cues during the simulation of route guidance.

3. The third block focused on the quality of the sound and how it was perceived by the user. The intention of this block was to determine if the user thought that the movement of the sound matched the audio cue which was triggered.

4. The next block tried to collect information regarding the users' expectations by addressing information regarding creating/updating/eliminating one of the spatial trajectories which were defined.

5. The last block compared the current solution offered by HERE Technologies and the proposed spatialisation of the audio cues.

Regarding the format of the answers presented to the user, the questions presented a pre-defined set of answers or an empty field to be filled. The usage of the 5 point scale together with the empty text field was determined as sufficient for the initial survey.

Blocks number 2 and 3 addressed the first RQ or the perception of the spatilisation. However, blocks 4 and 5 concentrated on the second one focusing on the user's expectations of the integration of the proposed feature.

Chapter 4

Findings

4.1 UX Expert Review

During the expert evaluation, several observations were made. In this section, those observations are divided into the previous phases mentioned in section 3.3.3. Between the phases, the observations were taken into consideration and applied the necessary changes in the next version of the prototype in order to expand its features. The observations are listed below:

4.1.1 First stage: Defining the trajectory of spatialisation

As mentioned above, during the first phase of the expert evaluation, observations regarding the spatialisation trajectories were made.

- Certain virtual sounds were rendered along linear trajectory ending on the line behind the user's head or the front one. The experts did not unequivocally perceive the direction or distance of these sounds. This situation was prone to cause front/rear auditory confusion, which is a typical problem in static binaural rendering due to ambiguous interaural cues and solely monaural spectral differences [9][21]

- During this first stage, certain trajectories reflected a different sound level while triggering the audio cue. Ideally, the sound level should remain constant during the spatialisation.

4.1.2 Second stage: Mechanics and perception

During the second stage, the experts focused on the perception of particular mechanics of the system and how the user could distinguish them.

- Upcoming trajectories were not described with coherence regarding the user's perceived real-world geometry. The need for circular trajectories was observed. Linear trajectories presented a good performance, but the localisation of the sound was not perceived accurately.
- Audio cues were overlapped in some driving situations, so the delivery to the final user was not clean. This observation was in the end, omitted. HERE's SDK applies speed-distance filters which avoid this overlapping. Therefore, the situation only appeared on the prototype and not to the final solution.
- During the testing some sound trajectories started not at the correct position creating some confusion while reproducing the audio cue.

4.1.3 Third stage: In-vehicle session

Finally, the third stage used the same scenario established in the second one: active route guidance. The difference between both stages was the environment where it was executed. In the second one, the navigation guidance was conducted as a digital simulation of commuting by car. Meanwhile, the third one was conveyed in a vehicle enabling a real scenario of driving. Therefore, the expert was able to follow the immersion methodology explained in section 3.3.3.2.

Once the three stages were conducted, a final version of the prototype was implemented. That version of the prototype was the one used during the user tests.

- During this phase, the experts noticed some missing trajectories to fulfil the spatial geometry of all the audio cues which were triggered during the test and therefore some new ones were defined.
- Parameters of the spatialisation were tweaked during the trip looking to provide a more realistic experience to the driver.
- Double turns were included in the prototype. When turns were not distancing much from each other, new audio cues were found. This audio cue guided the user to take the next two manoeuvres.
- Connection by Bluetooth directly the speakers of the vehicle was a success. The prototype could trigger the spatialisation.
- The connection was not closed when playing music on the radio. This was considered a milestone since it was once of the potential issues to be found.

- Minor UI/UX improvements were defined in this phase.

In the end, even if some minor thoughts were generated on particular trajectories, the prototype was able to spatialise the audio cues triggered by the SDK in real-time. The experts finally considered this functionality as a success since it enabled the spatialisation and implemented a new information source which was able to help the driver or the pedestrian to get a better understanding of the upcoming manoeuvre.

4.2 Survey

Once the online user test was conducted and the responses to the survey were gathered, the data was analysed. The test was conducted with 3 users, and the average completion time spent was 8 minutes and 33 seconds which was close the 10 minutes predicted during the design of the survey. The observations made by the users are summarized below:

Each of the users perceived that the sound was moving around them regardless of the audio system which was used during the test confirming that the spatial sound was correctly displayed. Indeed, users defined the context information (location of the next turn) delivered through the new spatial audio guidance as excellent.

> *"Really liked the moving voice, helped me a lot to focus and understand when I have to take my next turn and which direction it would be."*

Certainly, all of the users marked in the survey that there was no trajectory which could be defined as unnecessary and therefore should be eliminated. Nevertheless, while users believed that the perceived trajectories were coherent with what the voice was communicating, some concerns were noted regarding the spatialisation of one of the trajectories.

In general, users positively concluded that the new feature in question helped them to identify the direction of the next turn, and made it easier to remember it even when executing another task in parallel.

During the interview, one of the users also mentioned that the selection of the routes was really accurate since he felt that all the potential audio cues which he had in mind were covered. The reaction of the user made it possible to believe that the trajectories which were designed were suitable with the ones that the user had in mind.

> *"Felt good, seems that most driving cases are reflected by the selection and it felt quite natural."*

Overall, users rated the spatial audio guidance proposal presented a much better solution regarding the perceived information about the next turn direction in the proposed spatial audio solution. Still, two users mentioned that they were looking forward to seeing the functionality into more products of the company.

> *"It is way more contextual and gives better orientation and preparation for a turn."*
>
> *"It makes easier to remember the direction of the turn."*

After analysing the data gathered, a few observations were made. In general, the feature was welcomed positively but users expressed some expectations which were not fulfilled.

The observations which were made after analysing the data gathered through the survey:

- The users perceived audio movement during the guidance.
- One particular trajectory was identified as unusual. The users were expected to be prompted faster, to have a more realistic perception of the turn.
- The feature was welcome by the users. Indeed, they showed clear interest in future implementation.
- Integrating spatial audio helped the users not only to identify the direction the next turn but also helped them to keep it in mind when the next manoeuvre was triggered after a long time.
- Users notified the lack of one manoeuvre to be spatialised.

Chapter 5

Discussion

In this section of the thesis are mentioned the outcome obtained after analysing the data collected on both methods and crossing the results looking for similarities between the observations taken.

The information gathered during the UX expert review and the user test showed a correlation between them. Since experts' and users' perception of the sound movement was generally defined as a comprehensible, the results showed an acceptance of the proposed methodology used to define the spatial audio trajectories. The implications of this correlation establish a common set of expectations between the designers of the trajectories and the users.

Nevertheless, the analysis of the data suggested that, while the trajectories seemed to be perceived as coherent with what the voice communicates, it is required to evaluate different alternatives. Users particularly pointed out to the timing of spatialisation of one of the trajectories. Therefore, it was also detected a minor discrepancy between users' and experts' expectations. Nevertheless, this gap did not represent a conflict, rather than an increment of the validity of the defined trajectories, since users and experts agreed on moslty of the potential trajectories to be implemented.

Regarding the feasibility and benefits of the newly implemented feature, in line with the hypothesis generated back in the days by HERE UX team, the users showed satisfaction with the proposed functionality. Furthermore, both methods, expert review and user test showed an increment of awareness of the direction of the next turn on the listener.

Lastly, the sound movement around the car, rather a static mono audio cue, was more likely to catch the driver's attention when the audio guidance was not active recently. To be more specific, experts and users seemed to react better to an audio cue, for example, when driving for some kilometres in the highway and waiting for the next maneuver to take the next exit. This reaction could have occurred due to unfamiliarity the spatial feature in a navigation guidance system, or the expectation of identifying the next instruction, so it was also considered as a limitation.

Chapter 6

Conclusions

This research aimed to identify the feasibility of integrating a new spatial audio feature into HERE Technologies' SDK and evaluating if the resulting spatial trajectories designed during the process matched with the expectations the users. Based on the qualitative analysis of the data gathered during the expert review and the user test an acceptance of the proposed spatial audio feature can be concluded, making it possible to be considered as an important factor when navigating. In the future, the company also expect to find a connection of improvement of the safety, time-response and performance perceived by the driver when spatialising the cues triggered during the audio guidance designed for this thesis.

The results of the research indicated that drivers are more receptive to understand and localise the direction of the next turn when the audio cue is accompanied by audio movement. Still, the generalizability of the results is limited by the number of participants during the user test.

Based on these results, practitioners should also consider that the methodological choices were constrained by the social distance imposed in 2020 due to the pandemic of COVID-19, making it not possible to follow the original plan of execution. To better understand the implications of these results, future studies could address a new user test to gather quantitative data seeking to justify the validity of the defined spatial trajectories.

Concluding, looking at the first research question of *"How does the user perceive the condition of the spatialisation of the audiocues"*, the results were successful at detecting potential applications and advantages when integrating spatial audio in a navigation system. The experience obtained by users and experts was overall positive. The sound movement felt smooth so it was easy to identify the position of the source and helped to identify the direction of the next turn.

In regards to the second research question of *"How does the spatialisation* *of the audio cues fulfil the user expectations in regards to the perception of the spatial audio"*, while one of the trajectories was compromised by the users, they generally expressed satisfaction with respect to the trajectories which were defined. Therefore, while the proposed alternative trajectory by the users needs to be reviewed, the designed trajectories matched the expectations of the users who tested the prototype.

6.1 Limitations and future work

The prototype accomplished all the requirements that were set up during the ideation phase. Nevertheless, it presented some limitations:

- Just a prototype: The application did not offer all the features which are provided by HERE's SDK. The prototype lacked filters when triggering the audio cues depending on the current speed and the distance of the next manoeuvre. This filter would have avoided overlapping some of the audio cues, so the result would have been cleaner.
- Limited quantity of participants: Due to the social distance imposed by COVID-19, the number of testers was reduced. Still, reaching experts in the design field made it possible to gather quality information.
- All the users who tested the spatial audio feature were men. It would be interesting to understand if women perceive the trajectories in the same way.
- Online testing: Since the test was conducted online, it was not possible to observe how the users the reactions of the user during the test. Furthermore, it users stayed at home, which made it not possible to experience the spatial movement in a real scenario
- Front/back was not accurately perceived. Nevertheless, this was out of the topic marked for this thesis.
- While the prototype was also tested in the vehicle, the time invested using the speakers of the vehicle was limited compared with the headphones setup.
- Unspecific setup during user test: Since the user test was conducted online, it was not possible to determine a precise audio system to be used. On one hand, this fact could help us to test if the resulting audio depending on the earphones/headphones selected, was a positive

outcome. Nevertheless, it could not ensure a pre-tested experience to the user.

Once the agenda defined for the thesis was executed, a series of future steps were marked for future research taking into consideration the potential limitations observed on the methodology and results of the thesis. A more concise study will be conducted in order to decide if it is possible to implement the solution proposed in this thesis on the navigation system offered by HERE Technologies. Still, it should be noted that the proposal of this thesis has shown an active interest of the company. Therefore, the established next steps are minor improvements which could be defined out of the scope of this thesis.

This being said, some steps which could be taken into consideration are considering the potential trajectories proposed by the users, and to convey a second user test, recruiting more participants to increase the reliability of the results. Furthermore, a new method could be established in order to get quantitative data addressing the safety, reaction or completion time of the driver.

At last, a more concise test of the implemented feature should be conveyed while driving. Some differences between the experience perceived while wearing headphones or driving might emerge.

In the end, only a test with a large enough number of naive subjects would be able to determine if the spatial audio feature works generally.

Bibliography

[1] N. Bevan, J. Carter, J. Earthy, T. Geis, and S. Harker. New iso standards for usability, usability reports and usability measures. volume 9731, pages 268–278, 07 2016.

[2] S. S. Borojeni, L. Chuang, W. Heuten, and S. Boll. Assisting Drivers with Ambient Take-Over Requests in Highly Automated Driving. In *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - Automotive'UI 16*, pages 237–244, Ann Arbor, MI, USA, 2016. ACM Press.

[3] C.-C. Chang, J. Sodnik, and L. N. Boyle. Don't Speak and Drive: Cognitive Workload of In-Vehicle Speech Interactions. In *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications Adjunct - Automotive'UI 16*, pages 99–104, Ann Arbor, MI, USA, 2016. ACM Press.

[4] A. Cooper. *The Inmates Are Running the Asylum: Why High Tech Products Drive Us Crazy and How to Restore the Sanity (2nd Edition)*. Pearson Higher Education, 2004.

[5] P. J. Cooper and Y. Zheng. Turning gap acceptance decision-making: the impact of driver distraction. *Journal of Safety Research*, 33(3):321–335, Oct. 2002.

[6] R. F. Dam and T. Y. Siang. 5 Stages in the Design Thinking Process, 2020. Library Catalog: www.interaction-design.org.

[7] D. Damböck, T. Weissgerber, M. Kienle, and K. Bengler. Evaluation of a Contact Analog Head-Up Display for Highly Automated Driving. page 10, 2012.

[8] J. S. Dumas and J. C. Redish. *A Practical Guide to Usability Testing*. Intellect Books, GBR, 1st edition, 1999.

[9] M. Frank and F. Zotter. Simple reduction of front-back confusion in static binaural rendering. 03 2018.

[10] R. H. Gilkey and T. R. Anderson, editors. *Binaural and spatial hearing in real and virtual environments*. Lawrence Erlbaum Associates, Mahwah, N.J, 1997.

[11] J. I. Giménez-Nadal, J. A. Molina, and J. Velilla. Trends in Commuting Time of European Workers: A Cross-Country Analysis. page 40, 2020.

[12] F. Grani, S. Scheunig, and S. Trento. Method, apparatus and computer program product for spatial auditory cues. (US10477338B1), Nov. 2019. Library Catalog: Google Patents.

[13] R. Gudjonsdottir. *Personas and scenarios in use*. PhD thesis, June 2010.

[14] w. James. *The Principles of Psychology*. Harvard University Press, 1981. Google-Books-ID: JKeCfOnHSA8C.

[15] P. W. Jordan. *Designing Pleasurable Products: An Introduction to the New Human Factors*. CRC Press, Aug. 2002. Google-Books-ID: 0s3el8sDjHsC.

[16] R. L. Klatzky, J. R. Marston, N. A. Giudice, R. G. Golledge, and J. M. Loomis. Cognitive load of navigating without vision when guided by virtual sound versus spatial language. *Journal of Experimental Psychology: Applied*, 12(4):223–232, 2006.

[17] A. J. Kolarik, B. C. J. Moore, P. Zahorik, S. Cirstea, and S. Pardhan. Auditory distance perception in humans: a review of cues, development, neuronal bases, and effects of sensory loss. *Attention, Perception & Psychophysics*, 78:373–395, 2016.

[18] A. Kun, T. Paek, and Z. Medenica. The Effect of Speech Interface Accuracy on Driving Performance. page 5, 2007.

[19] C. Lallemand, V. Koenig, and G. Gronier. How relevant is an expert evaluation of user experience based on a psychological needs-driven approach? In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction Fun, Fast, Foundational - NordiCHI '14*, pages 11–20, Helsinki, Finland, 2014. ACM Press.

[20] Long and Frank. Real or Imaginary: The effectiveness of using personas in product design - Frontend. Jan. 2009.

[21] J. C. Makous and J. C. Middlebrooks. *Twodimensional sound localization by human listeners*, volume 85. 1990.

[22] D. G. Malham. Approaches to spatialisation. *Organised Sound*, 3(2):167–177, Aug. 1998.

[23] D. G. Malham and A. Myatt. 3-D Sound Spatialization using Ambisonic Techniques. *Computer Music Journal*, 19(4):58, 1995.

[24] J. C. Middlebrooks and D. M. Green. Sound Localization by Human Listeners. *Annual Review of Psychology*, 42(1):135–159, Jan. 1991.

[25] J. Nielsen. *Usability Engineering*. Morgan Kaufmann, Oct. 1994. Google-Books-ID: 95As2OF67f0C.

[26] D. R. Perrott, T. Sadralodabai, K. Saberi, and T. Z. Strybel. Aurally aided visual search in the central visual field: Effects of visual load and visual enhancement of the target. *Human Factors*, 33(4):389–400, 1991. Place: US Publisher: Human Factors & Ergonomics Society.

[27] V. Pulkki. Virtual Sound Source Positioning Using Vector Base Amplitude Panning. *Journal of the Audio Engineering Society*, 45(6):456–466, June 1997. Publisher: Audio Engineering Society.

[28] V. Pulkki. Localization of Amplitude-Panned Virtual Sources II: Two- and Three-Dimensional Panning | Request PDF, 2001. Library Catalog: www.researchgate.net.

[29] R. Rabenstein and S. Spors. Spatial Sound Reproduction with Wave Field Synthesis. volume 105, Jan. 2005.

[30] D. A. Redelmeier and R. J. Tibshirani. Association between cellular-telephone calls and motor vehicle collisions. *The New England Journal of Medicine*, 336(7):453–458, Feb. 1997.

[31] A. S. Masters of sound, 2017. Library Catalog: newsroom.porsche.com.

[32] J. Sodnik, C. Dicke, S. Tomažič, and M. Billinghurst. A user study of auditory versus visual interfaces for use while driving. *International Journal of Human-Computer Studies*, 66(5):318–332, May 2008.

[33] D. L. Strayer, F. A. Drews, and D. J. Crouch. A Comparison of the Cell Phone Driver and the Drunk Driver. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 48(2):381–391, June 2006.

[34] M. T. The sound of the BMW Concept i4., 2020. Library Catalog: www.press.bmwgroup.com.

[35] I. Tashev, M. Seltzer, Y.-C. Ju, Y.-Y. Wang, and A. Acero. Commute UX: Voice Enabled In-car Infotainment System. page 7, 2009.

[36] A. Vetek and S. Lemmelä. Could a dialog save your life?: analyzing the effects of speech interaction strategies while driving. In *Proceedings of the 13th international conference on multimodal interfaces - ICMI '11*, page 145, Alicante, Spain, 2011. ACM Press.

[37] C. D. W. Processing Resources in Attention, Dual Task Performance, and Workload Assessment. page 61, 1981.

[38] P. Zahorik. Assessing auditory distance perception using virtual acoustics. *The Journal of the Acoustical Society of America*, 111(4):1832–1846, Apr. 2002.

[39] E. Zoller. Key vendor rankings and market trends: June 2019 update. page 23, 2019.

Appendix A

Appendix methods

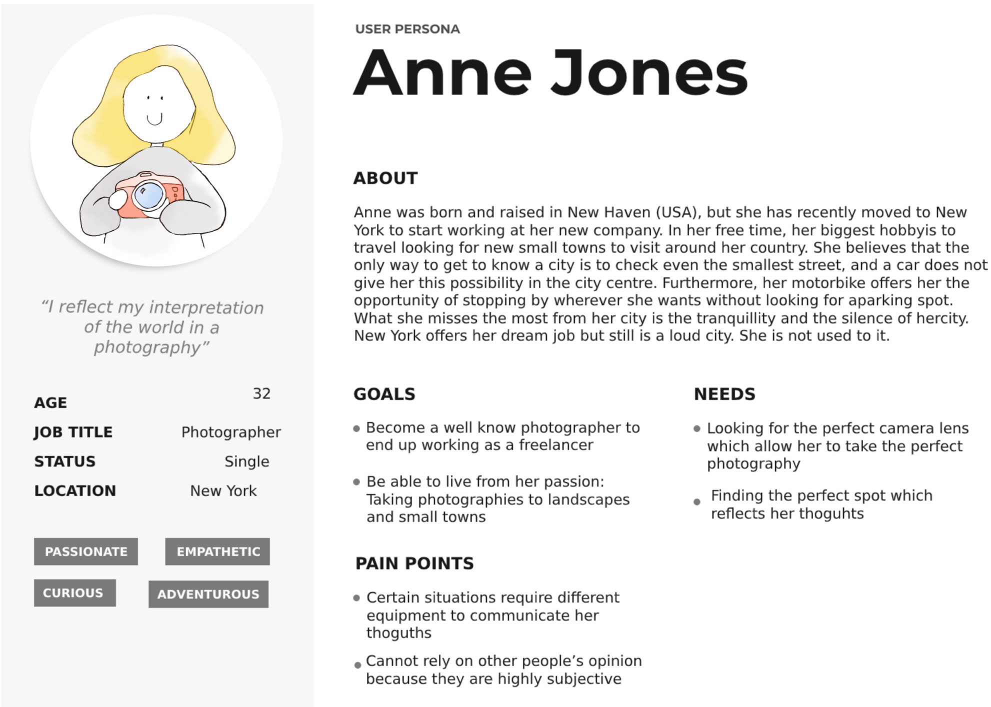


Figure A.1: Layout persona

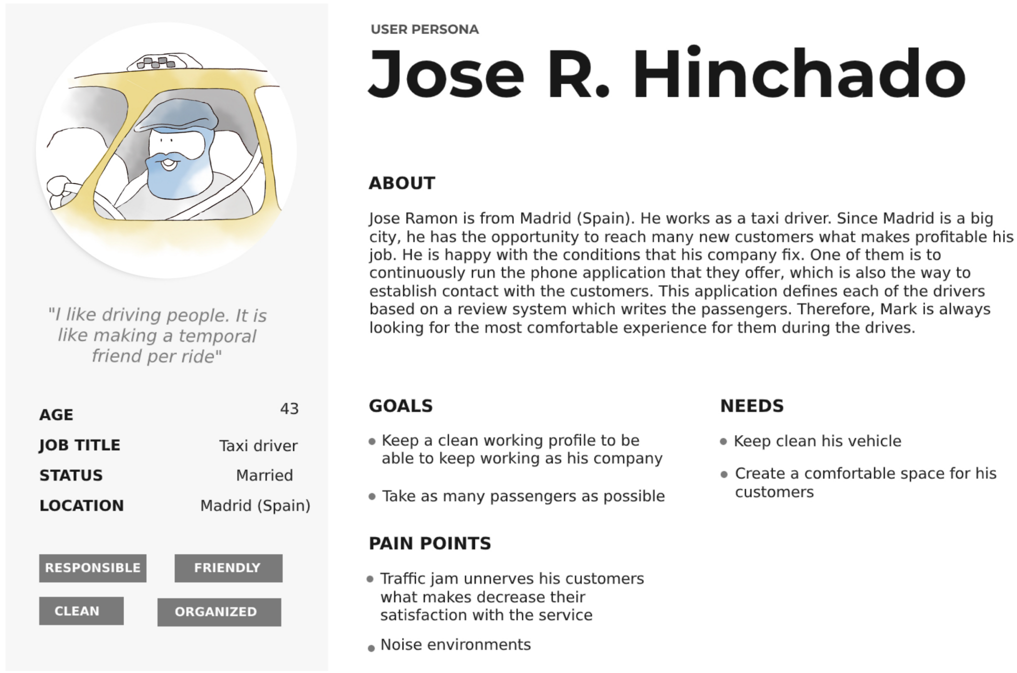


Figure A.2: Layout persona

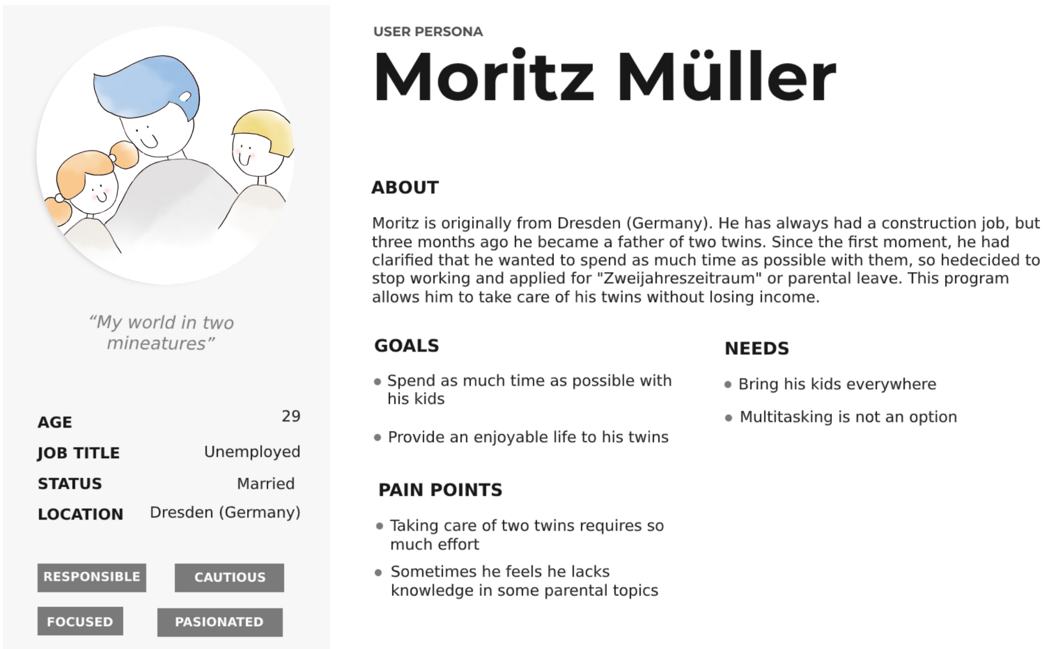


Figure A.3: Layout persona

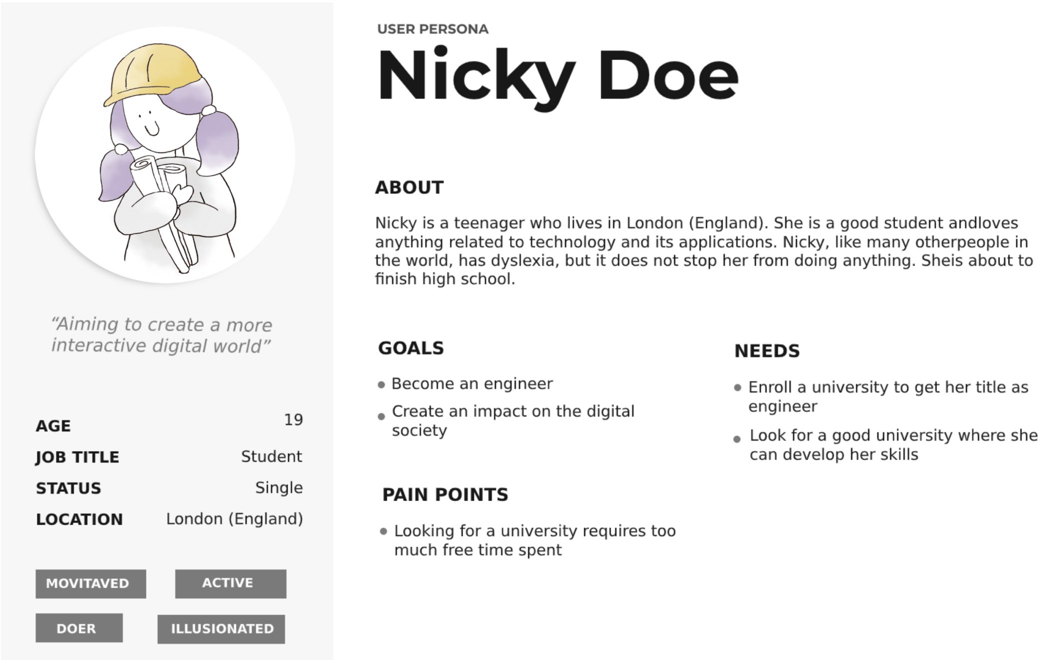


Figure A.4: Layout persona